DT01 Rec'd PCT/FT0 12 3 DEC 2004

Sequence listing

<120> DNA sequence and preparation of grass pollen allergen Phl p 4 by recombinant methods

Sequenz-Protokoll

```
<110> Merck Patent GmbH
<120> DNA-Sequenz und rekombinante Herstellung des Graspollen-Allergens
Phl p 4
<130> P 02/101
<140> EP 02 013953.1
<141> 2002-06-25
<160> 52
<170> PatentIn version 3.1
<210> 1
<211> 1503
<212> DNA
<213>
      Phleum pratense
<220>
<221> artificial_DNA_sequence
<222>
       (1)..(69)
<223>
      DNA sequence derived from sequenced protein
<220>
<221> native_DNA_sequence
<222>
       (70)..(1503)
<223>
```

<220>

<221> CDS <222> (1)..(1503) <223> <400> 1 ' tac ttc ccg ccg ccg gct gct aaa gaa gac ttc ctg ggt tgc ctg gtt 48 Tyr Phe Pro Pro Pro Ala Ala Lys Glu Asp Phe Leu Gly Cys Leu Val aaa gaa ato ccg ccg cgt ctg ttg tac gcg aaa tcg tcg ccg gcg tat Lys Glu Ile Pro Pro Arg Leu Leu Tyr Ala Lys Ser Ser Pro Ala Tyr ccc tca gtc ctg ggg cag acc atc cgg aac tcg cgg tgg tcg tcg ccg Pro Ser Val Leu Gly Gln Thr Ile Arg Asn Ser Arg Trp Ser Ser Pro 144 40 gac aac gtg aag ccg atc tac atc gtc acc ccc acc aac gcc tcc cac 192 Asp Asn Val Lys Pro Ile Tyr Ile Val Thr Pro Thr Asn Ala Ser His atc cag tcc gcc gtg gtg tgc ggc cgc cgg cac ggt gtc cgc atc cgc Ile Gln Ser Ala Val Val Cys Gly Arg Arg His Gly Val Arg Ile Arg 240 gtg cgc agc ggc ggg cac gac tac gag ggc ctc tcg tac cgg tcc ctg Val Arg Ser Gly Gly His Asp Tyr Glu Gly Leu Ser Tyr Arg Ser Leu 288 cag ccc gag gag ttc gcc gtc gtc gac ctt agc aag atg cgg gcc gtg 336 Gin Pro Glu Glu Phe Ala Val Val Asp Leu Ser Lys Met Arg Ala Val tgg gtg gac ggg aag gcc cgc acg gcg tgg gtc gac tcc ggc gcg cag Trp Val Asp Gly Lys Ala Arg Thr Ala Trp Val Asp Ser Gly Ala Gln 384 120 ctc ggc gag ctc tac tac gcc atc cac aag gcg agt aca gtg ctg gcg 432 Leu Gly Glu Leu Tyr Tyr Ala Ile His Lys Ala Ser Thr Val Leu Ala 130 ttc ccg gcc ggc gtg tgc ccg acc atc ggc gtg ggc ggc aac ttc gcg Phe Pro Ala Gly Val Cys Pro Thr Ile Gly Val Gly Asn Phe Ala 480 ggc ggc ggc ttc ggc atg ctg ctg cgc aag tac ggc atc gcg gcc gag Gly Gly Phe Gly Met Leu Leu Arg Lys Tyr Gly Ile Ala Ala Glu 528 aac gtc atc gac gtg aag ctc gtc gac gcc aac ggc acg ctg cac gac Asn Val Ile Asp Val Lys Leu Val Asp Ala Asn Gly Thr Leu His Asp 576 aag aag too atg ggc gac gac cat tto tgg gcc gtc agg ggc ggc ggg Lys Lys Ser Met Gly Asp Asp His Phe Trp Ala Val Arg Gly Gly Gly 624 205

•																		
	ggc Gly	gag Glu 210	agc Ser	ttc Phe	ggc Gly	atc Ile	gtg Val 215	gtc Val	gcg Ala	tgg Trp	aag Lys	gtg Val 220	agg Arg	ctc Leu	ctg Leu	ccg Pro		672
	gtg Val 225	ccg Pro	ccc Pro	acg Thr	gtg Val	acc Thr 230	gtg Val	ttc Phe	aag Lys	atc Ile	ccc Pro 235	aag Lys	aag Lys	gcg Ala	agc Ser	gag Glu 240		720
	ggc Gly	gcc Ala	.gtg Val	gac Asp	atc Ile 245	atc Ile	aac Asn	agg Arg	tgg Trp	cag Gln 250	gtg Val	gtc Val	gcg Ala	ccg Pro	cag Gln 255	ctc Leu		768
	ccc Pro	gac Asp	gac Asp	ctc Leu 260	atg Met	atc Ile	cgc Arg	gtc Val	atc Ile 265	gcg Ala	cag Gln	ggc Gly	ccc Pro	acg Thr 270	gcc Ala	acg Thr		816
	ttc Phe	gag Glu	gcc Ala 275	atg Met	tac Tyr	ctg Leu	ggc Gly	acc Thr 280	tgc Cys	caa Gln	acc Thr	ctg Leu	acg Thr 285	ccg Pro	atg Met	atg Met		864
	agc Ser	agc Ser 290	aag Lys	ttc Phe	ccg Pro	gag Glu	ctc Leu 295	ggc Gly	atg Met	aac Asn	gcc Ala	tcg Ser 300	cac His	tgc Cys	aac Asn	gag Glu		912
	atg Met 305	tcg Ser	tgg Trp	atc Ile	cag Gln	tcc Ser 310	atc Ile	ccc Pro	ttc Phe	gtc Val	cac His 315	ctc Leu	ggc Gly	cac His	agg Arg	gac Asp 320		960
	aac Asn	atc Ile	gag Glu	gac Asp	gac Asp 325	ctc Leu	ctc Leu	aac Asn	cgg Arg	aac Asn 330	aac Asn	acc Thr	ttc Phe	aag Lys	CCC Pro 335	ttc Phe		1008
	gcc Ala	gaa Glu	tac Tyr	aag Lys 340	tcg Ser	gac Asp	tac Tyr	gtc Val	tac Tyr 345	gag Glu	ccg Pro	ttc Phe	ccc Pro	aag Lys 350	agg Arg	gtg Val		1056
	tgg Trp	gag Glu	cag Gln 355	atc Ile	ttc Phe	agc Ser	acc Thr	tgg Trp 360	ctc Leu	ctg Leu	aag Lys	ccc Pro	ggc Gly 365	gcg Ala	Gly	atc Ile		1104
															gag Glu			1152
															cag Gln			1200
	gtc Val	aac Asn	tac Tyr	tgg Trp	ttc Phe 405	gcc Ala	ccg Pro	gga Gly	gcc Ala	ggc Gly 410	gcg Ala	gcg Ala	cca Pro	ttg Leu	tcg Ser 415	tgg Trp		1248
	agc Ser	aag Lys	gag Glu	atc Ile 420	tac Tyr	aac Asn	tac Tyr	atg Met	gag Glu 425	cca Pro	tac Tyr	gtg Val	agc Ser	aag Lys 430	aac Asn	ccc Pro	.:·	1296
	agg Arg	cag Gln	gcc Ala 435	tac Tyr	gcc Ala	aac Asn	tac Tyr	agg Arg 440	gac Asp	atc Ile	gac Asp	ctc Leu	ggg Gly 445	agg Arg	aac Asn	gag Glu	•	1344
	gtg	gtg	aac	gac	gtc	tcc	acc	ttc	agc	agc	ggt	ttg	gtg	tgg	ggc _.	cag		1392

Val Val Asn Asp 450	Val Ser Thr 455	Phe Ser Ser Gly	Leu Val Trp 0	Gly Gln · ·
aaa tac ttc aag Lys Tyr Phe Lys 465	ggc aat ttc Gly Asn Phe 470	cag agg ctc gcc Gln Arg Leu Ala 475	Ile Thr Lys C	ggc aag 1440 Gly Lys 480
gtg gat ccc acc Val Asp Pro Thr	gac tac ttc Asp Tyr Phe 485	agg aac gag cag Arg Asn Glu Gln 490	Ser Ile Pro E	ccg ctc 1488 Pro Leu 195
atc aaa aag tac Ile Lys Lys Tyr 500	tga		·	. 1503
<210> 2				•
<211> 500				
<212> PRT				
<213> Phleum pr	atense			
<400> 2				
Tyr Phe Pro Pro 1	Pro Ala Ala 5 .	Lys Glu Asp Phe		Leu Val 15
Lys Glu Ile Pro 20	Pro Arg Leu	Leu Tyr Ala Lys 25	Ser Ser Pro A	Ala Tyr
Pro Ser Val Leu 35	Gly Gln Thr	Ile Arg Asn Ser 40	Arg Trp Ser S	Ser Pro
Asp Asn Val Lys 50	Pro Ile Tyr 55	Ile Val Thr Pro	Thr Asn Ala S	Ser His
Ile Gln Ser Ala 65	Val Val Cys 70	Gly Arg Arg His	Gly Val Arg I	Ile Arg 80
Val Arg Ser Gly	Gly His Asp	Tyr Glu Gly Leu 90		Ser Leu 95
Gln Pro Glu Glu 100	Phe Ala Val	Val Asp Leu Ser 105	Lys Met Arg 1	Ala Val
Trp Val Asp Gly 115	Lys Ala Arg	120	Asp Ser Gly 1	Ala Gln
Leu Gly Glu Leu 130	Tyr Tyr Ala 135		Ser Thr Val 1	Leu Ala

Phe 145	Pro	Ala	Gly	Val	Cys 150	Pro	Thr	Ile	Gly	Val 155	Gly	Gly	Asn	Phe	Ala 160
Gly	G1 y	Gly	Phe	Gly 165	Met	Leu	Leu	Arg	Lys 170	Tyr	Gly	Ile	Ala	Ala 175	Glu
Asn	Val	Ile	Asp 180	Val	Lys	Leu	Val	Asp 185	Ala	Asn	Gly	Thr	Leu 190	His	Asp
Lys	Lys	Ser 195	Met	Gly	Asp	Asp	His 200	Phe	Trp	Ala	Val	Arg 205	Gly	Gly	Gly
Gly	Glu 210	Ser	Phe	Gly	Ile	Val 215	Val	Ala	Trp	Lys	Val 220	Arg	Leu	Leu	Pro
Val 225	Pro	Pro	Thr	Val	Thr 230	Val	Phe	Lys	Ile	Pro 235	Lys	Lys	Ala	Ser	Glu 240
Gly	Ala	Val	Asp	11e 245		Asn	Arg	Trp	G1n 250	Val	Val	Ala	Pro	Gln 255	Leu
Pro	Asp	Asp	Leu 260	Met	Ile	Arg	Val	Ile 265	Ala	Gln	Gly	Pro	Thr 270	Ala	Thr
Phe	Glu	Ala 275	Met	Tyr	Leu	Gly	Thr 280	Суз	Gln	Thr	Leu	Thr 285	Pro	Met	Met
Ser	Ser 290	Lys	Phe	Pro	Glu	Leu 295	Gly	Met	Asn	Ala	Ser 300	His	Cys	Asn	Glu
Met 305	Ser	Trp	Ile	Glń	Ser 310	Ile	Pro	Phe	Val	His 315	Leu	GĴĄ	His	Arg	Asp 320
Asn	Ile	Glu	Asp	Asp 325	Leu	Leu	Asn	Arg	Asn 330	Asn	Thr	Phe	Lys	Pro 335	Phe
Ala	Glu	Tyr	Lys 340	Ser	Asp	Tyr	Val	Tyr 345	Glu	Pro	Phe	Pro	Lys 350	Arg	Val
Trp	Glu	Gln 355	Ile	Phe	Ser	Thr	Trp 360	Leu	Leu	Lys	Pro	Gly 365	Ala	Gly	Ile
Met	Ile 370	Phe	Asp	Pro	Tyr	Gly 375	Ala	Thr	Ile	Ser	Ala 380	Thr	Pro	Glu	Trp

```
Ala Thr Pro Phe Pro His Arg Lys Gly Val Leu Phe Asn lle Gln Tyr
Val Asn Tyr Trp Phe Ala Pro Gly Ala Gly Ala Ala Pro Leu Ser Trp
                                    410
Ser Lys Glu Ile Tyr Asn Tyr Met Glu Pro Tyr Val Ser Lys Asn Pro
Arg Gln Ala Tyr Ala Asn Tyr Arg Asp Ile Asp Leu Gly Arg Asn Glu
Val Val Asn Asp Val Ser Thr Phe Ser Ser Gly Leu Val Trp Gly Gln
Lys Tyr Phe Lys Gly Asn Phe Gln Arg Leu Ala Ile Thr Lys Gly Lys
Val Asp Pro Thr Asp Tyr Phe Arg Asn Glu Gln Ser Ile Pro Pro Leu
Ile Lys Lys Tyr
<210> 3
<211> 1503
<212>
      DNA
<213>
       Phleum pratense
<220>
<221>
       artificial_DNA_sequence
```

<222> (1)..(69)

DNA sequence derived from sequenced protein <223>

<220>

<221> native_DNA_sequence

<222> (70)..(1503)

<223>

- 7 -

· <220>															
<221> CDS	<221> CDS														
<222> (1)(15	<222> (1)(1503)														
<223>	<223>														
<400> 3	g ccg gct gct aa	a.gaa gac ttc	cta gat tac c	tg gtt 48											
	Pro Ala Ala Ly			eu Val											
	g ccg cgt ctg tt p Pro Arg Leu Le														
	g ggg cag acc at n Gly Gln Thr Il 40	e Arg Asn Ser													
	g ccg atc tac at s Pro Ile Tyr Il 55														
atc cag tcc gcc Ile Gln Ser Ala 65	gtg gtg tgc gg Nal Val Cys Gl 70	c cgc cgg cac y Arg Arg His 75	ggt gtc cgc at Gly Val Arg I	tc cgc . 240 le Arg 80											
	ggg cac gac ta Gly His Asp Ty 85			er Leu											
	g ttc gcc gtc gt g Phe Ala Val Va g														
tgg gtg gac ggg Trp Val Asp Gly 115	, aag gee ege ae , Lys Ala Arg Th 12	r Ala Trp Val	gac tcc ggc gg Asp Ser Gly Al 125	cg cag 384 la Gln											
	tac tac gcc at Tyr Tyr Ala Il 135														
ttc ccg gcc ggc Phe Pro Ala Gly 145	gtg tgc ccg ac Val Cys Pro Th 150	c atc ggc gtg r Ile Gly Val 155	ggc ggc aac tt Gly Gly Asn Pl	tc gcg 480 ne Ala 160											
	ggc atg ctg ct Gly Met Leu Le 165		Gly Ile Ala A												
	: gtg aag ctc gt) Val Lys Leu Va)														
	ggc gac gac ca Gly Asp Asp Hi 20	s Phe Trp Ala													

ggc Gly	gag Glu 210	agc Ser	ttc Phe	ggc Gly	atc Ile	gtg Val 215	gtc Val	gcg Ala	tgg Trp	aag Lys	gtg Val 220	agg Arg	ctc Leu	ctg Leu	ccg Pro	672
gtg Val 225	ccg Pro	ccc Pro	acg Thr	gtg Val	acc Thr 230	gtg Val	ttc Phe	aag Lys	atc Ile	ccc Pro 235	aag Lys	aag Lys	gcġ Ala	agc Ser	gag Glu 240	720
ggc Gly	gcc Ala	gtg Val	gac Asp	atc Ile 245	atc Ile	aac Asn	agg Arg	tgg Trp	cag Gln 250	gtg Val	gtc Val	gcg Ala	ccg Pro	cag Gln 255	ctc Leu	768
ccc Pro	gac Asp	gac Asp	ctc Leu 260	atg Met	atc Ile	cgc Arg	gtc Val	atc Ile 265	gcg Ala	cag Gln	ggc Gly	ccc Pro	acg Thr 270	gcc Ala	acg Thr	816
ttc Phe	gag Glu	gcc Ala 275	atg Met	tac Tyr	ctg Leu	ggc Gly	acc Thr 280	tgc Cys	caa Gln	acc Thr	ctg Leu	acg Thr 285	ccg Pro	atg Met	atg Met	864
agc Ser	agc Ser 290	aag Lys	ttc Phe	ccc Pro	gag Glu	ctc Leu 295	G1Å ååc	atg Met	aac Asn	gcc Ala	tcg Ser 300	cac His	tgc Cya	aac Asn	gag Glu	912
atg Met 305	tcg Ser	tgg Trp	atc Ile	cag Gln	tcc Ser 310	atc Ile	ccc Pro	ttc Phe	gtc Val	cac His 315	ctc Leu	ggc Gly	cac His	agg Arg	gac Asp 320	960
aac Asn	atc Ile	gag Glu	gac Asp	gac Asp 325	ctc Leu	ctc Leu	aac Asn	cgg Arg	aac Asn 330	aac Asn	acc Thr	ttc Phe	aag Lys	ccc Pro 335	ttc Phe	1008
gcc Ala	gaa Glu	tac Tyr	aag Lys 340	tcg Ser	gac Asp	tac Tyr	gtc Val	tac Tyr 345	gag Glu	ccg Pro	ttc Phe	ccc Pro	aag Lys 350	gaa Glu	gtg Val	1056
tgg Trp	gag Glu	cag Gln 355	atc Ile	ttc Phe	agc Ser	acc Thr	tgg Trp 360	ctc Leu	ctg Leu	aag Lys	ccc Pro	ggc Gly 365	gcg Ala	ggg Gly	atc Ile	1104
atg Met	atc Ile 370	ttc Phe	gac Asp	ccc Pro	tac Tyr	ggc Gly 375	gcc Ala	acc Thr	atc Ile	agc Ser	gcc Ala 380	acc Thr	ccg Pro	gag Glu	tgg Trp	1152
gcg Ala 385	acg Thr	ccg Pro	ttc Phe	cct Pro	cac His 390	cgc Arg	aag Lys	ggc Gly	gtc Val	ctc Leu 395	ttc Phe	aac Asn	atc Ile	cag Gln	tac Tyr 400	. 1200
gtc Val	aac Asn	tac Tyr	tgg Trp	ttc Phe 405	gcc Ala	ccg Pro	gga Gly	gcc Ala	ggc Gly 410	gcg Ala	·gcg Ala	cca Pro	ttg Leu	tcg Ser 415	tgg	1248
agc Ser	aag Lys	gag Glu	atc Ile 420	Tyr	aac Asn	tac Tyr	atg Met	gag Glu 425	cca Pro	tac Tyr	gtg Val	agc Ser	aag Lys 430	aac Asn	.ccc Pro	1296
agg Arg	cag Gln	gcc Ala 435	tac Tyr	gcc Ala	aac Asn	tac Tyr	agg Arg 440	gac Asp	atc Ile	gac Asp	ctc Leu	999 Gly 445	agg. Arg	aac Asn	gag Glu	1344
gtg	gtg	aac	gac	gtc	tcc	acc	ttc	agc	agc	ggt	ttg	gtg	tgg	ggc	cag	1392

-9-

Val Val Asn Asp Val Ser Thr Phe Ser Ser Gly Leu Val Trp Gly Gln 455 aaa tac ttc aag ggc aat ttc cag agg ctc gcc atc acc aag ggc aag 1440 Lys Tyr Phe Lys Gly Asn Phe Gln Arg Leu Ala Ile Thr Lys Gly Lys 470 gtg gat ccc acc gac tac ttc agg aac gag cag agc atc ccg ctc Val Asp Pro Thr Asp Tyr Phe Arg Asn Glu Gln Ser Ile Pro Pro Leu 1488 490 1503 atc aaa aag tac tga Ile Lys Lys Tyr 500 <210> <211> <212> PRT Phleum pratense <400> 4 Tyr Phe Pro Pro Pro Ala Ala Lys Glu Asp Phe Leu Gly Cys Leu Val Lys Glu Ile Pro Pro Arg Leu Leu Tyr Ala Lys Ser Ser Pro Ala Tyr Pro Ser Val Leu Gly Gln Thr Ile Arg Asn Ser Arg Trp Ser Ser Pro Asp Asn Val Lys Pro Ile Tyr Ile Val Thr Pro Thr Asn Ala Ser His 50 55 60 Ile Gln Ser Ala Val Val Cys Gly Arg Arg His Gly Val Arg Ile Arg Val Arg Ser Gly Gly His Asp Tyr Glu Gly Leu Ser Tyr Arg Ser Leu Gln Pro Glu Glu Phe Ala Val Val Asp Leu Ser Lys Met Arg Ala Val Trp Val Asp Gly Lys Ala Arg Thr Ala Trp Val Asp Ser Gly Ala Gln Leu Gly Glu Leu Tyr Tyr Ala Ile His Lys Ala Ser Pro Val Leu Ala 135 140

							•								
Phe 145	Pro	Ala	Gly	Val	Cys 150	Pro	Thr	Ile	Gly	Val 155	Gly	Gly	Asn	Phe	Ala 160
Gly	Gly	Gly	Phe	Gly 165	Met	Leu	Leu	Arg	Lys 170	Tyr	Gly	Ile	Ala	Ala 175	Glu
Asn	Val	Ile	Asp 180	Val	Lys	Leu	Val	Asp 185	Ala	Asn	Gly	Thr	Leu 190	His	Asp
Lys	Lys	Ser 195	Met	Gly	Asp	Asp	His 200	Phe	Trp	Ala	Val	Arg 205	Gly	Gly	Gly
Gly	Glu 210	Ser	Phe	Gly	Ile	Val 215	Vål	Ala	Trp	Lys	Val 220	Arg	Leu	Leu	Pro
Val 225	Pro	Pro	Thr	Val	Thr 230	Val	Phe	Lys	Ile	Pro 235	Lys	Lys	Ala	Ser	Glu 240
Gly	Ala	Val	Asp	Ile 245	Ile	Asn	Arg	Trp	Gln 250	Val	Val	Ala	Pro	Gln 255	Leu
Pro	Asp	Asp	Leu 260	Met	Ile	Arg	Val	Ile 265	Ala	Gln	Gly	Pro	Thr 270	Ala	Thr
Phe	Glu	Ala 275	Met	Tyr	Leu	Gly	Thr 280	Cys	Gln	Thr	Leu	Thr 285	Pro	Met	Met
Ser	Ser 290	Lys	Phe	Pro	Glu	Leu 295	Gly	Met	Asn	Ala	Ser 300	His	Cys	Asn	Glu
Met 305	Ser	Trp	Ile	Gln	Ser 310	Ile	Pro	Phe	Val	His 315	Leu	Gly	His	Arg	Asp 320
Asn	Ile	Glu	Asp	Asp 325	Leu	Leu	Asn	Arg	Asn 330	Asn	Thr	Phe	Lys	Pro 335	Phe
Ala	Glu	Tyr	Lys 340	Ser	Asp	Tyr		Tyr 345		Pro	Phe	Pro	Lys 350		Val
Trp	Glu	Gln 355	Ile	Phe	Ser	Thr	Trp 360	Leu	Leu	Lys	Pro	Gly 365	Ala	Gly	Ile
Met	Ile 370	Phe	Asp	Pro	Tyr	Gly 375	Ala	Thr	Ile	Ser	Ala 380	Thr	Pro	Glu	Trp

```
Ala Thr Pro Phe Pro His Arg Lys Gly Val Leu Phe Asn Ile Gln Tyr
Val Asn Tyr Trp Phe Ala Pro Gly Ala Gly Ala Ala Pro Leu Ser Trp
Ser Lys Glu Ile Tyr Asn Tyr Met Glu Pro Tyr Val Ser Lys Asn Pro
            420
Arg Gln Ala Tyr Ala Asn Tyr Arg Asp Ile Asp Leu Gly Arg Asn Glu 435 440 445
Val Val Asn Asp Val Ser Thr Phe Ser Ser Gly Leu Val Trp Gly Gln
Lys Tyr Phe Lys Gly Asn Phe Gln Arg Leu Ala Ile Thr Lys Gly Lys
Val Asp Pro Thr Asp Tyr Phe Arg Asn Glu Gln Ser Ile Pro Pro Leu
Ile Lys Lys Tyr
<210> 5
<211>
      1503
<212>
<213>
      Phleum pratense
<220>
      artificial_DNA_sequence
<221>
<222>
       (1)..(69)
      DNA sequence derived from sequenced protein
<223>
<220>
```

native_DNA_sequence

(70) .. (1503)

<222>

<223>

PCT/EP03/06092

- 12 -

<220> <221> CDS (1)..(1503) <222> <400> 5 tac ttc ccg ccg ccg gct gct aaa gaa gac ttc ctg ggt tgc ctg gtt Tyr Phe Pro Pro Pro Ala Ala Lys Glu Asp Phe Leu Gly Cys Leu Val 96 aaa qaa atc ccg ccg cgt ctg ttg tac gcg aaa tcg tcg ccg gcg tat Lys Glu Ile Pro Pro Arg Leu Leu Tyr Ala Lys Ser Ser Pro Ala Tyr ccc tca gtc ctg ggg cag acc atc cgg aac tcg agg tgg tcg tcg ccg Pro Ser Val Leu Gly Gln Thr Ile Arg Asn Ser Arg Trp Ser Ser Pro gac aac gtg aag ccg ctc tac atc atc acc ccc acc aac gtc tcc cac Asp Asn Val Lys Pro Leu Tyr Ile Ile Thr Pro Thr Asn Val Ser His 192 atc cag tcc gcc gtg gtg tgc ggc cgc cgc cac agc gtc cgc atc cgc Ile Gln Ser Ala Val Val Cys Gly Arg Arg His Ser Val Arg Ile Arg 240 gtg cgc agc ggc ggg cac gac tac gag ggc ctc tcg tac cgg tct ttg Val Arg Ser Gly Gly His Asp Tyr Glu Gly Leu Ser Tyr Arg Ser Leu 288 cag ccc gag acg ttc gcc gtc gtc gac ctc aac aag atg cgg gcg gtg Gln Pro Glu Thr Phe Ala Val Val Asp Leu Asn Lys Met Arg Ala Val 336 tgg gtg gac ggc aag gcc cgc acg gcg tgg gtg gac tcc ggc gcg cag Trp Val Asp Gly Lys Ala Arg Thr Ala Trp Val Asp Ser Gly Ala Gln 384 ctc ggc gag ctc tac tac gcc atc tat aag gcg agc ccc acg ctg gcg Leu Gly Glu Leu Tyr Tyr Ala Ile Tyr Lys Ala Ser Pro Thr Leu Ala 432 ttc ccg gcc ggc gtg tgc ccg acg atc gga gtg ggc ggc aac ttc gcg Phe Pro Ala Gly Val Cys Pro Thr Ile Gly Val Gly Gly Asn Phe Ala 480 ggc ggc ggc ttc ggc atg ctg ctg cgc aag tac ggc atc gcc gcg gag Gly Gly Gly Phe Gly Met Leu Leu Arg Lys Tyr Gly Ile Ala Ala Glu 528 aac gtc atc gac gtg aag ctc gtc gac gcc aac ggc aag ctg cac gac 576 Asn Val Ile Asp Val Lys Leu Val Asp Ala Asn Gly Lys Leu His Asp aag aag too atg ggc gac gac cat tto tgg gcc gtc agg ggc ggc ggg Lys Lys Ser Met Gly Asp Asp His Phe Trp Ala Val Arg Gly Gly Gly 624

ggc Gly	gag Glu 210	agc Ser	ttc Phe	ggc Gly	atc Ile	gtg Val 215	gtc Val	gcg Ala	tgg Trp	cag Gln	gtg Val 220	aag Lys	ctc Leu	ctg Leu	ccg Pro	672
					aca Thr 230											720
					atc Ile											768
					atc Ile											816
ttc Phe	gag Glu	gcc Ala 275	atg Met	tac Tyr	ctc Leu	ggc Gly	acc Thr 280	tgc Cys	aaa Lys	acc Thr	ctg Leu	acg Thr 285	ccg Pro	ttg Leu	atg Met	864
					gag Glu											912
					tcc Ser 310											960
					ctc Leu											1008
					gac Asp											1056
					aac Asn											1104
					tac Tyr											1152
					cac His 390											. 1200
gtc Val	aac Asn	tac Tyr	tgg Trp	ttc Phe 405	gcc Ala	ccg Pro	gga Gly	gcc Ala	gcc Ala 410	gcc Ala	gcg Ala	ccc Pro	ctc Leu	tcg Ser 415	tgg Trp	1248
					aac Asn											1296
					aac Asn											1344
gtg	gtc	aac.	gac	gtc	tcc	acc	tac	gcc	agc	ggc	aag	gtc	tgg	ggc	cag	1392/

Val Val Asn 450	Asp Val S	er Thr Tyr 455	Ala Ser Gly	Lys Val Trp 460	Gly Gln
	Lys Gly A			att acc aag Ile Thr Lys	
gtc gat cct Val Asp Pro	acc gac t Thr Asp T 485	ac ttc agg yr Phe Arg	aac gag cag Asn Glu Gln 490	agc atc ccg Ser Ile Pro	ccg ctc 1488 Pro Leu 495
atc aaa aag Ile Lys Lys	_				1503
<210> 6					•
<211> 500					
<212> PRT					·
<213> Phle	m pratens				•
					•
<400> 6			•		
Tyr Phe Pro	Pro Pro A 5	la Ala Lys	Glu Asp Phe 10	Leu Gly Cys	Leu Val 15
Lys Glu Ile	Pro Pro A 20	rg Leu Leu	Tyr Ala Lys 25	Ser Ser Pro 30	Ala Tyr
Pro Ser Val	Leu Gly G	ln Thr Ile 40	Arg Asn Ser	Arg Trp Ser	Ser Pro
Asp Asn Val 50	Lys Pro L	eu Tyr Ile 55	Ile Thr Pro	Thr Asn Val	Ser His
Ile Gln Ser 65	Ala Val V		Arg Arg His	Ser Val Arg	Ile Arg
Val Arg Ser	Gly Gly H 85	is Asp Tyr	Glu Gly Leu 90 .	Ser Tyr Arg	Ser Leu
Gln Pro Glu	Thr Phe A	la Val Val	Asp Leu Asn 105	Lys Met Arg 110	Ala Val
Trp Val Asp	Gly Lys A	la Arg Thr 120		Asp Ser Gly 125	Ala Gln
Leu Gly Glu 130	Leu Tyr T	yr Ala Ile 135	Tyr Lys Ala	Ser Pro Thr 140	Leu Ala

Phe Pro Ala Gly Val Cys Pro Thr Ile Gly Val Gly Gly Asn Phe Ala Gly Gly Gly Phe Gly Met Leu Leu Arg Lys Tyr Gly Ile Ala Ala Glu Asn Val Ile Asp Val Lys Leu Val Asp Ala Asn Gly Lys Leu His Asp Lys Lys Ser Met Gly Asp Asp His Phe Trp Ala Val Arg Gly Gly Gly Gly Glu Ser Phe Gly Ile Val Val Ala Trp Gln Val Lys Leu Leu Pro Val Pro Pro Thr Val Thr Ile Phe Lys Ile Ser Lys Thr Val Ser Glu Gly Ala Val Asp Ile Ile Asn Lys Trp Gln Val Val Ala Pro Gln Leu Pro Ala Asp Leu Met Ile Arg Ile Ile Ala Gln Gly Pro Lys Ala Thr 265 Phe Glu Ala Met Tyr Leu Gly Thr Cys Lys Thr Leu Thr Pro Leu Met Ser Ser Lys Phe Pro Glu Leu Gly Met Asn Pro Ser His Cys Asn Glu 295 300 Met Ser Trp Ile Gln Ser Ile Pro Phe Val His Leu Gly His Arg Asp Ala Leu Glu Asp Asp Leu Leu Asn Arg Asn Asn Ser Phe Lys Pro Phe Ala Glu Tyr Lys Ser Asp Tyr Val Tyr Gln Pro Phe Pro Lys Thr Val Trp Glu Gln Ile Leu Asn Thr Trp Leu Val Lys Pro Gly Ala Gly Ile Met Ile Phe Asp Pro Tyr Gly Ala Thr Ile Ser Ala Thr Pro Glu Ser

Ala Thr Pro Phe Pro His Arg Lys Gly Val Leu Phe Asn Ile Gln Tyr 385 390 395 400

Val Asn Tyr Trp Phe Ala Pro Gly Ala Ala Ala Ala Pro Leu Ser Trp
405 410 415

Ser Lys Asp Ile Tyr Asn Tyr Met Glu Pro Tyr Val Ser Lys Asn Pro 420 425 430

Arg Gln Ala Tyr Ala Asn Tyr Arg Asp Ile Asp Leu Gly Arg Asn Glu 435 440 445

Val Val Asn Asp Val Ser Thr Tyr Ala Ser Gly Lys Val Trp Gly Gln 450 455 460

Lys Tyr Phe Lys Gly Asn Phe Glu Arg Leu Ala Ile Thr Lys Gly Lys 465 470 475 480

Val Asp Pro Thr Asp Tyr Phe Arg Asn Glu Gln Ser Ile Pro Pro Leu 485 490 495

Ile Lys Lys Tyr 500

<210> 7

<211> 10

<212> PRT

<213> Phleum pratense

<220>

<221> MISC_FEATURE

<222> (6)..(6)

<223> undetermined amino acid

<400> 7

Ile Val Ala Leu Pro Xaa Gly Met Leu Lys

<210> 8

<211> 14

```
<212> PRT
<213> Lolium perenne
<400> 8
Phe Leu Glu Pro Val Leu Gly Leu Ile Phe Pro Ala Gly Val
<210>
<211>
<212> PRT
<213> Lolium perenne
<400> 9
Gly Leu Ile Glu Phe Pro Ala Gly Val
<210> 10
<211> 12
<212> PRT
<213> Dactylus glomerata
<400> 10
Asp Ile Tyr Asn Tyr Met Glu Pro Tyr Val Ser Lys
<210> 11
<211> 11
<212> PRT
<213> Dactylus glomerata
<400> 11
Val Asp Pro Thr Asp Tyr Phe Gly Asn Glu Gln 1 . 5 10
<210> 12
```

<400> 15

```
<211> 17
<212>
       Dactylus glomerata
<400> 12
Ala Arg Thr Ala Trp Val Asp Ser Gly Ala Gln Leu Gly Glu Leu Ser
1 10 15
Tyr
<210> 13
<211> 15
<212> PRT
<213> Dactylus glomerata
<400> 13-
Gly Val Leu Phe Asn Ile Gln Tyr Val Asn Tyr Trp Phe Ala Pro
<210> 14
<211> 11
<212> PRT
<213> Cynodon dactylon '
<400> 14
Lys Thr Val Lys Pro Leu Tyr Ile Ile Thr Pro
<210> 15
<211> 22
<212> PRT
<213> Cynodon dactylon
```

```
Lys Gln Val Glu Arg Asp Phe Leu Thr Ser Leu Thr Lys Asp Ile Pro 1 5 10 15
Gln Leu Tyr Leu Lys Ser
<210> 16
<211> 16
<212> PRT
<213> Cynodon dactylon
<400> 16
Thr Val Lys Pro Leu Tyr Ile Ile Thr Pro Ile Thr Ala Ala Met Ile
                                      10
<210> 17
<211> 24
<212> PRT
<213> Cynodon dactylon
<400> 17
Leu Arg Lys Tyr Gly Thr Ala Ala Asp Asn Val Ile Asp Ala Lys Val
Val Asp Ala Gln Gly Arg Leu Leu
<210> 18
<211> 14
<212> PRT
<213> Cynodon dactylon
<400> 18
Lys Trp Gln Thr Val Ala Pro Ala Leu Pro Asp Pro Asn Met 1 \phantom{0} 10
<210> 19
```

```
<211> 15
<212> PRT
<213> Cynodon dactylon
<400> 19
Val Thr Trp Ile Glu Ser Val Pro Tyr Ile Pro Met Gly Asp Lys 1 5 10 15
<210>
      20
<211> 19
<212> PRT
<213> Cynodon dactylon
<220>
<221> MISC_FEATURE
<222> (8)..(8)
<223> undetermined amino acid
<400> 20
Gly Thr Val Arg Gln Leu Leu Xaa Arg Thr Ser Asn Ile Lys Ala Phe : 10 15
Gly Lys Tyr .
<210> 21
<211>, 23
<212> PRT
<213> Cynodon dactylon
<400> 21
Thr Ser Asn Ile Lys Ala Phe Gly Lys Tyr Lys Ser Asp Tyr Val Leu 1 5 10 15
Glu Pro Ile Pro Lys Lys Ser
```

```
<210> 22
<211>
       13
      PRT
<212>
<213> Cynodon dactylon
<400> 22
Tyr Arg Asp Leu Asp Leu Gly Val Asn Gln Val Val Gly
<210> 23
<211> 15
<212> PRT
<213> Cynodon dactylon
<400> 23
Ser Ala Thr Pro Pro Thr His Arg Ser Gly Val Leu Phe Asn Ile
<210> 24
<211>
       36
<212> PRT
<213> Cynodon dactylon ·
<400> 24
Ala Ala Ala Leu Pro Thr Gln Val Thr Arg Asp Ile Tyr Ala Phe
Met Thr Pro Tyr Val Ser Lys Asn Pro Arg Gln Ala Tyr Val Asn Tyr 20 25 30
Arg Asp Leu Asp
<210> 25
<211> 149
```

~2127	DIVA						
<213>	Phle	eum pratense	e		_		
<400> caccgga	25 aagg	gggtgctgtt	caacatccag	tacgtcaact	actggttcgc	cccgggagcc	60
ggcgcgç	gege	cattgtcgtg	gagcaaggag	atctacaact	acatggagcc	gtacgtgagc	- 120
aaggaco	ccg	tccaggccta	cgccaacta				149
<210>	26						
<21i>	299						
<212>	DNA						
<213>	Phle	eum pratense	e			•	
<400> actacto	26 ggtt	cgccccggga	gccggcgcgg	cgccattgtc	gtggagcaag	gagatctaca	60
actacat	gga	gccatacgtg	agcaagaacc	ccaggcaggc	ctacgccaac	tacagggaca	120
tegacet	cegg	gaggaacgag	gtggtgaacg	acgtctccac	cttcagcagc	ggtttggtgt	180
ggggcca	agaa	atacttcaag	ggcaacttcc	agaggctcgc	catcaccaag	ggcaäggtgg	240
atccca	ccga	ctacttcagg	aacgagcaga	gcatcccgcc	gctcatcaaa	aagtactga	299
<210>	27						
<211>	33						
<212>	PRT		•	-			
<213>	Phle	eum pratens	е		· •		
					. *		
<220>				•			•
<221>	MISC	_FEATURE					
<222>	(14))(14)	ī	•			
<223>	und	etermined a	mino acid				
<400>	27	•					
Tyr Ph	e Pr	o Pro Pro A	la Ala Lys	Glu Asp Phe	Leu Gly Xa	a Leu Val	

```
Lys Glu Ile Pro Pro Arg Leu Leu Tyr Ala Lys Ser Ser Pro Ala Tyr 20 25 30
Pro '
<210>
<211> 18
<212> PRT
<213> Phleum pratense
<220>
<221> MISC_FEATURE
<222> (6)..(6)
<223> undetermined amino acid
<400> 28
Ser Ala Thr Pro Phe Xaa His Arg Lys Gly Val Leu Phe Asn Ile Gln
Tyr Val
<210> 29
<211>
<212> PRT
<213> Phleum pratense
<220>
<221> MISC_FEATURE
<222> (3)..(8)
<223> undetermined amino acid
<400> 29
Gly Leu Xaa Tyr Arg Xaa Leu Xaa Pro Glu
```

```
<210>
       30
       12
<211>
<212>
<213>
       Phleum pratense
<220>
<221> MISC_FEATURE
<222>
      (2)..(9)
       undetermined amino acid
<223>
<400> 30
Lys Xaa Met Gly Asp Asp His Phe Xaa Ala Val Arg
1 5 10
<210>
       31
<211>
<212> PRT
<213> Phleum pratense
<400> 31
Ala Pro Glu Gly Ala Val Asp Ile Ile 1
<210>
       32
<211>
       16
       PRT
<212>
<213> Phleum pratense
<400> 32
Met Glu Pro Tyr Val Ser Ile Asn Pro Val Gln Ala Tyr Ala Asn Tyr
<210> 33
```

<400>

```
<211> 15
<212> PRT
<213> Phleum pratense
<220>
<221> MISC_FEATURE
<222> (14)..(14)
<223> undetermined amino acid
<400> 33
Tyr Phe Pro Pro Pro Ala Ala Lys Glu Asp Phe Leu Gly Xaa Leu
<210> 34
<211> 10
<212> PRT
<213> Phleum pratense
<400> 34
Leu Tyr Ala Lys Ser Ser Pro Ala Tyr Pro
<210>
      35
<211>
      33
<212> PRT
<213> Phleum pratense
<220>
<221> MISC_FEATURE
<222>
      (14)..(14)
<223> undetermined amino acid
```

```
Tyr Phe Pro Pro Pro Ala Ala Lys Glu Asp Phe Leu Gly Xaa Leu Val
Lys Glu Ile Pro Pro Arg Leu Leu Tyr Ala Lys Ser Ser Pro Ala Tyr \cdot 20 25 30
Pro
      36
<210>
<211>
      29
<212> PRT
<213> Phleum pratense
<220>
<221> MISC_FEATURE
<222> (14)..(14)
<223> undetermined amino acid
<400> 36
Tyr Phe Pro Pro Pro Ala Ala Lys Glu Asp Phe Leu Gly Xaa Leu Val
Lys Glu Pro Pro Arg Leu Leu Tyr Ala Lys Ser Ser Pro
<210> 37
<211> 15 -
<212> PRT
<213> Phleum pratense
<220>
<221> MISC_FEATURE
<222>
      (4)..(14)
      undetermined amino acid
```

```
<400> 37
Tyr Phe Pro Xaa Xaa Ala Ala Lys Glu Asp Phe Leu Gly Xaa Leu
<210> 38
<211> 15
<212> PRT
<213> Phleum pratense
<220>
<221> MISC_FEATURE
<222>
      (4)..(14)
<223> undetermined amino acid
<400> 38
Tyr Phe Pro Xaa Xaa Ala Lys Lys Glu Asp Phe Leu Gly Xaa Leu
<210> 39
<211> 15
<212> PRT
<213> Phleum pratense
<220>
<221> MISC_FEATURE
<222> (4)..(14)
      undetermined amino acid
<223>
<400> .39
Tyr Phe Pro Xaa Xaa Ala Ala Lys Asp Asp Phe Leu Gly Xaa Leu
<210> 40
<211> 11
```

```
<212> PRT
<213> Phleum pratense
<220>
<221> MISC_FEATURE
<222>
      (4)..(5)
<223> undetermined amino acid
<400> 40
Tyr Phe Pro Xaa Xaa Leu Ala Asn Glu Asp Phe 1 5 10
<210>
       41
<211> 18
<212> PRT
<213> Phleum pratense
<220>
<221> MISC_FEATURE
<222>
      (6)..(6)
<223>
      undetermined amino acid
<400> 41
Ser Ala Thr Pro Phe Xaa His Arg Lys Gly Val Leu Phe Asn Ile Gln
Tyr Val
<210> 42
<211>
<212> PRT
<213> Phleum pratense
```

```
<220>
<221> MISC_FEATURE
<222> (3)..(8)
<223> undetermined amino acid
<400> 42
Gly Leu Xaa Tyr Arg Xaa Leu Xaa Pro Glu
<210> 43
<211> 12
<212> PRT
<213> Phleum pratense
<220>
<221> MISC_FEATURE
<222> (2)..(9)
<223> undetermined amino acid
<400> 43
<210> 44
<211> 9
<212> PRT
<213> Phleum pratense.
<400> 44
Ala Pro Glu Gly Ala Val Asp Ile Ile . . . 1
<210> 45
<211> 16
```

<210> 48

```
<212> PRT
<213> Phleum pratense
<400> 45
Met Glu Pro Tyr Val Ser Ile Asn Pro Val Gln Ala Tyr Ala Asn Tyr
                                    10
<210> 46
<211> 29
<212> DNA
<213> Phleum pratense
<220>
<221> misc_feature
<222>
      (1)..(29)
      'n' means inosin
<223>
<400> 46
ytntaygcna arwsnwsncc ngcntaycc
                                                                       29
<210> 47
<211>
      28
<212> DNA
<213> Phleum pratense
<220>
<221> misc_feature
      (1)..(28)
<222>
<223>
       'n' means inosin
<400> 47
caymgnaarg gngtnytntt yaayatmc
                                                                       28
```

<211>	26		
<212>	DNA		
<213>	Phleum pratense		
<220>			
<221>	misc_feature		
<222>	(1)(26)		
<223>	'n' means inosin		
<400>	48 gcrt angcytgnac nggrtt		26
carcen	gore angeresme nggree		
<210>	49		•
<211>	23		
<212>	DNA		
<213>	Phleum pratense		
	•		
<400>	49 ggtt cgccccggga gcc	•	23
<210>	50		
<211>	28		
<212>	DNA		
<213>	Phleum pratense	•	
<400> tgaagt	50 attt ctggccccac accaaacc		28
<210>	51		
<211>	24		
<212>	DNA		
<213>	Phleum pratense		
<400> cccttg	51 gtga tggcgagcot ctgg		24

- 32 -

<210> 52

<211> 23

<212> DNA

<213> Phleum pratense

<400> 52 ctcagtcctg gggcagacca tcc